**OPEN**

Correspondence and
requests for materials
should be addressed to
X.-P.M. (miaoxp@mail.
hust.edu.cn) or F.-Q.G.
(gongfangqi63@
gmail.com)

* These authors
contributed equally to
this work.

# The Roles of Ca²⁺/NFAT Signaling Genes in Kawasaki Disease: Single- and Multiple-Risk Genetic Variants

Wei Wang[1]*, Jiao Lou[2]*, Rong Zhong[2], Yan-qi Qi[1], Na Shen[2], Xu-zai Lu[3], Yu-jia Wang[4], Qing Zhang[1], Li Zou[2], Jia-yu Duan[2], Jun-tao Ke[2], Xiao-ping Miao[2] & Fang-qi Gong[1]

[1]Children's Hospital, Zhejiang University School of Medicine, Hangzhou, PR China, [2]Department of Epidemiology and Biostatistics and Key Laboratory of Environment and Health, Ministry of Education & Ministry of Environmental Protection, and State Key Laboratory of Environmental Health (Incubating), School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, PR China, [3]Guangdong Women and Children Hospital, Guangzhou, PR China, [4]Centre Hospitalier de l'Université de Montréal, CRCHUM-Hôpital Notre-Dame, Pavillion DeSève, Montreal, Canada.

Ca²⁺/nuclear factor of activated T-cells (Ca²⁺/NFAT) signaling pathway may play a crucial role in Kawasaki disease (KD). We investigated 16 genetic variants, selected by bioinformatics analyses or previous studies, in 7 key genes involved in this pathway in a Chinese population. We observed a significantly or marginally increased KD risk associated with rs2720378 GC + CC genotypes (OR = 1.39, 95% CI = 1.07–1.80, $P$ = 0.014) or rs2069762 AC + CC genotypes (OR = 1.28, 95% CI = 0.98–1.67, $P$ = 0.066), compared with their wild type counterparts. In classification and regression tree analysis, individuals carrying the combined genotypes of rs2720378 GC or CC genotype, rs2069762 CA or CC genotype and rs1561876 AA genotype exhibited the highest KD risk (OR = 2.12, 95% CI = 1.46–3.07, $P$ < 0.001), compared with the lowest risk carriers of rs2720378 GG genotype. Moreover, a significant dose effect was observed among these three variants ($P_{trend}$ < 0.001). In conclusion, this study implicates that single- and multiple-risk genetic variants in this pathway might contribute to KD susceptibility. Further studies on more comprehensive single nucleotide polymorphisms, different ethnicities and larger sample sizes are warranted, and the exact biological mechanisms need to be further clarified.

K awasaki disease (KD; OMIM 611775), also called mucocutaneous lymph node syndrome, is an acute, self-limited vasculitis that predominantly affects children younger than 5 years old, which was first reported by Tomisaku Kawasaki from Japan in 1974 in the English language literature[1]. KD occurs worldwide and is mostly prevalent in East Asian population, such as Japanese[2], Koreans[3] and Taiwanese[4]. The major clinical manifestations of KD include prolonged fever, bilateral non-purulent conjunctivitis, diffuse mucosal inflammation, polymorphous skin rashes, peripheral extremity changes, and cervical lymphadenopathy[5–6]. Approximately 15–25% of untreated and 3–5% of treated children develop coronary artery lesions, including coronary artery dilatation or coronary artery aneurysm, and even myocardial infarction[7,8], making KD the leading cause of acquired heart disease in childhood in the developed countries.

Although almost 40 years have passed since the first description of the disease, its etiology is still not completely clear. However, the familial aggregation of KD patients suggests that genetic factors contribute to KD susceptibility and outcome[9,10]. Previous association studies based on the candidate gene approach have identified a set of common variants contributing to KD risk, but few received consistent replications. *Inositol 1,4,5-triphosphate 3-kinase C* (*ITPKC*) gene may be one of the most studied targets in genetic susceptibility of KD. A genome-wide linkage analysis conducted in Japanese KD sibling-pair samples led to the identification of a functional single nucleotide polymorphism (SNP), rs28493229 in the *ITPKC* gene on chromosome 19q13.2[11,12], which was confirmed subsequently by a meta-analysis integrating 10 case-control studies and 2 transmission/disequilibrium tests[13]. *ITPKC* encodes one of the three isoenzymes of inositol 1,4,5-triphosphate 3-kinase that phosphorylates inositol 1,4,5-triphosphate, and is involved in the Ca²⁺/nuclear factor of activated T-cells (NFAT) signaling pathway in T cells as a negative regulator[12]. Besides, another SNP rs2290692 in this gene was also found associated with KD susceptibility in a Chinese population[14].

It is believed that T-cell activation plays an important role in the pathogenesis of vascular endothelial cell injury by eliciting proinflammatory reactions at the onset of KD[15,16], among which, Ca²⁺/NFAT signaling pathway may

1

**Table 1 | Association between individual SNP and KD risk**

| Chr | Gene | rs ID | MAF in CHB* | W/H/V frequency | | Reference/ Variant allele | OR† | 95% CI | LR P‡ |
|-----|------|-------|-------------|-----------------|---|---------------------------|-----|--------|-------|
| | | | | KD cases | Controls | | | | |
| 4 | | | | | | | | | |
| | CASP3 | rs2720378 | 0.326 | 184/186/56 | 253/191/49 | G/C | **1.39**§ | **1.07–1.80** | **0.014** |
| | IL-2 | rs2069762 | 0.239 | 170/193/55 | 230/210/52 | A/C | 1.28§ | 0.98–1.67 | 0.066 |
| | STIM2 | rs10263 | 0.367 | 174/196/53 | 207/220/65 | A/G | 0.94¶ | 0.64–1.39 | 0.760 |
| 11 | | | | | | | | | |
| | STIM1 | rs1561876 | 0.344 | 244/151/23 | 266/193/32 | A/G | 0.84§ | 0.65–1.10 | 0.204 |
| | STIM1 | rs3750994 | 0.244 | 298/112/16 | 334/137/20 | T/G | 0.91§ | 0.69–1.21 | 0.529 |
| | STIM1 | rs3750996 | 0.133 | 249/142/19 | 287/172/25 | A/G | 0.94§ | 0.72–1.23 | 0.663 |
| 18 | | | | | | | | | |
| | NFATc1 | rs9966033 | 0.378 | 158/209/53 | 184/222/74 | T/C | 0.79¶ | 0.54–1.16 | 0.230 |
| | NFATc1 | rs754093 | 0.311 | 176/197/45 | 222/206/63 | T/G | 0.82¶ | 0.55–1.23 | 0.338 |
| | NFATc1 | rs7227107 | 0.367 | 167/201/54 | 194/224/68 | A/G | 0.90¶ | 0.62–1.33 | 0.600 |
| 19 | | | | | | | | | |
| | ITPKC | rs28493229 | 0.070‖ | 359/63/4 | 422/71/0 | G/C | 1.11§ | 0.77–1.59 | 0.573 |
| | ITPKC | rs10420685 | 0.178 | 230/158/38 | 267/184/40 | A/G | 1.10§ | 0.69–1.76 | 0.675 |
| | ITPKC | rs2290692 | 0.433 | 109/190/112 | 111/238/119 | G/C | 0.86§ | 0.64–1.17 | 0.344 |
| | ITPKC | rs2561530 | 0.270‖ | 230/160/33 | 258/202/28 | A/G | 1.39¶ | 0.83–2.34 | 0.215 |
| | ITPKC | rs4802085 | 0.500 | 129/187/103 | 129/238/102 | G/C | 1.17¶ | 0.86–1.60 | 0.317 |
| 20 | | | | | | | | | |
| | NFATc2 | rs6013193 | 0.330 | 150/210/67 | 157/251/84 | T/G | 0.87§ | 0.66–1.14 | 0.302 |

Abbreviations: MAF, minor allele frequency; CHB, Han Chinese in Beijing; Chr, chromosome; W, wild type homozygote; H, heterozygote; V, variant homozygote; OR, odds ratio; 95%CI, 95% confidence interval; LR, logistic regression
*MAF was downloaded from the online database of HapMap for Han Chinese in Beijing, China
†OR calculation was conducted under assumption that variant alleles were risk alleles
‡All the P values were adjusted for gender. The positive results were in bold
§The inheritance model of the risk alleles was dominant
¶The inheritance model of the risk alleles was recessive
‖MAF was downloaded from the online database of 1000 Genomes for Han Chinese in Beijing, China

play a crucial role. Once the T-cell receptor receives a stimulus, phospholipase Cγ1 is activated, and then generates amount of diacylglycerol and inositol 1,4,5-triphosphate, the latter of which then binds to its receptor expressed on endoplasmic reticulum (ER) membrane and causes the release of $Ca^{2+}$ into the cytoplasm. The depletion of $Ca^{2+}$ store in ER leads to a process that extracellular $Ca^{2+}$ enters through calcium released-activated $Ca^{2+}$ channels on the plasma membrane, which is evoked by stromal interaction molecule (STIM) as a sensor of $Ca^{2+}$ in ER. Calcineurin is activated by the bond between cytoplasmic $Ca^{2+}$ and calmodulin, afterwards dephosphorylates NFAT in the cytoplasm and leads nuclear translocation of NFAT. NFAT in the nucleus drives transcription of genes important in T cell activation such as *IL-2*[17].

Caspase 3 (CASP3) is one of the effector caspases that play an important role in the execution phase of apoptosis. It has been reported that CASP3 cleaves NFATc2 as one of its substrates in T cells[18], and also acts as a negative regulator of the $Ca^{2+}$/NFAT pathway. In addition, another study conducted by Onouchi et al. revealed that a G to A substitution of one commonly associated SNP located in the 5′-untranslated region of *CASP3* (rs113420705, formerly rs72689236) abolished binding of NFAT to the DNA sequence surrounding the SNP[19].

Compared to *ITPKC* and *CASP3*, few studies have investigated on other loci involved in the $Ca^{2+}$/NFAT signaling pathway and KD susceptibility, such as *STIM1*, *STIM2*, *NFATc1*, *NFATc2*, *IL-2*, in spite of their contributions to the immune and inflammatory system[20–25]. Therefore, we proposed our hypothesis that multiple common genetic variants involved in $Ca^{2+}$/NFAT signaling pathway may individually or interactively contribute to the etiology and pathogenesis of KD. Referring to the interactions, it could be challenging for the traditional analytic strategies, such as logistic regression (LR), to fully characterize them since sparseness of the data in high dimensions would lead to large standard errors of parameter estimates, and

result in an increase in type I errors[26,27]. What's more, statistical power would decrease and type II errors would increase when detecting interactions by LR in a relatively small sample size[28]. Thus, a nonparametric data mining approach, classification and regression tree (CART) has been applied to explore high-order gene-gene and gene-environment interactions, given its statistical advantages in overcoming the inaccuracy parameter estimates and providing good power for identifying high-order interactions[29].

Here, we investigated 16 SNPs which were predicted to be functional or identified by previous studies in 7 candidate genes involved in the $Ca^{2+}$/NFAT signaling pathway in a case-control study of a Chinese population. We examined the individual and combined effects of these 16 variants by traditional unconditional LR model and high-order gene-gene interactions in modulating KD risk using CART analysis.

## Results

**Subjects characteristics.** A total of 428 KD children and 493 healthy controls were enrolled in this study. There were 263 (61%) males in cases and 303 (61%) males in controls, and no statistically significant difference was observed between cases and controls in the distribution of gender (Pearson $\chi^2 = 0.000$, $P = 0.997$).

**Association analysis between individual SNP and KD risk.** The genotyping call rates of all the 16 SNPs were >95% except CASP3 rs113420705 (92%), so that the SNP rs113420705 was not analyzed any further. The genotype distributions of the remaining 15 SNPs in our control subjects were all in the Hardy-Weinberg equilibrium (HWE, $P > 0.05$) and their minor allele frequencies (MAFs) were similar to those in HapMap database of Han Chinese in Beijing, China (CHB, Supplementary Table S1 online). As shown in Table 1, two SNPs, *CASP3* rs2720378 and *IL-2* rs2069762 were signifi-
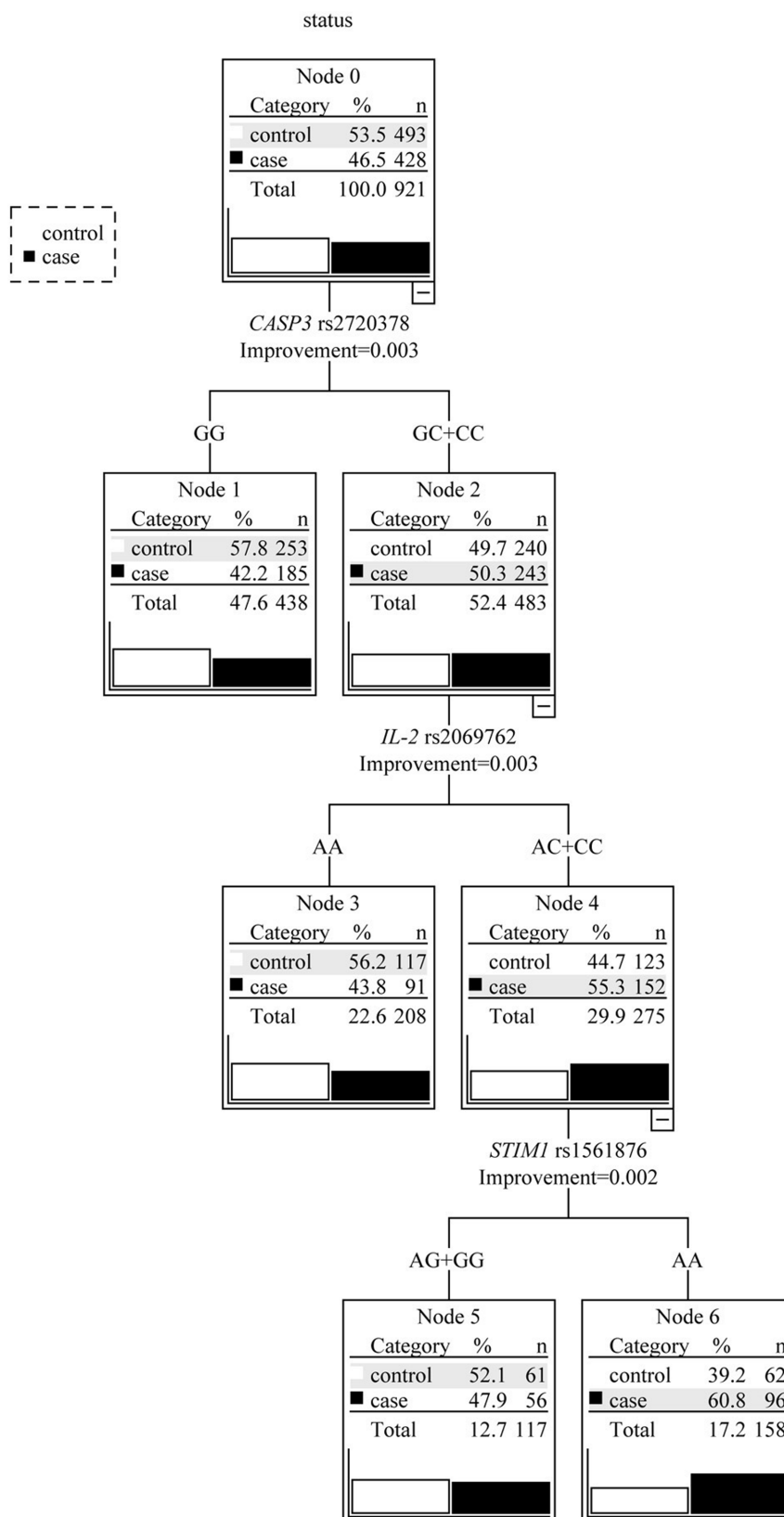
**Figure 1 | CART analysis of genetic variants in Ca²⁺/NFAT signaling pathway and KD risk.** Cases and controls are denoted by white and black box respectively, and each node contains frequencies and percentages of cases and controls in each subgroup.

cantly or marginally associated with the increased risk of KD (odds ratio (OR) = 1.39, 95% confidence interval (CI) = 1.07–1.80, $P$ = 0.014; OR = 1.28, 95% CI = 0.98–1.67, $P$ = 0.066), both under the dominant model.

**Association of high-order interactions with KD risk by CART analysis.** In the final optimal decision tree generated by the CART analysis (Figure 1), the initial split of the root node was CASP3 rs2720378, and patients harboring rs2720378 C allele (GC or CC

**Table 2 | Risk estimates of CART terminal nodes**

| Node | Genotype | KD cases | Controls | Case rate* (%) | OR† | 95% CI | LR P‡ |
|---|---|---|---|---|---|---|---|
| 1 | rs2720378 (W) | 185 (43.22) | 253 (51.32) | 42.2 | Ref. | Ref. | |
| 3 | rs2720378 (HV)-rs2069762 (W) | 91 (21.26) | 117 (23.73) | 43.8 | 1.06 | 0.76–1.48 | 0.718 |
| 5 | rs2720378 (HV)-rs2069762 (HV)-rs1561876 (HV) | 56 (13.08) | 61 (12.37) | 47.9 | 1.26 | 0.83–1.89 | 0.276 |
| 6 | rs2720378 (HV)-rs2069762 (HV)-rs1561876 (W) | 96 (22.43) | 62 (12.58) | 60.8 | **2.12** | **1.46–3.07** | **<0.001** |

Abbreviations: CART, classification and regression tree; KD, Kawasaki disease; OR, odds ratio; 95%CI, 95% confidence interval; LR, logistic regression; W, wild type homozygote; H, heterozygote; V, variant homozygote
*Case rate is the percentage of KD patients among all individuals in each node
†ORs of terminal nodes were calculated by LR analysis adjusted for gender
‡All the P values were adjusted for gender. The positive results were in bold

genotype) had a higher risk to suffer KD compared with patients with GG genotype, suggesting that CASP3 rs2720378 was the strongest risk factor for KD among the 15 SNPs examined. A deeper exploration of the classification tree structure demonstrated distinct interaction patterns between individuals carrying rs2720378 GC or CC genotype and those with rs2720378 GG genotype. Individuals harboring GG genotype had the lowest risk for KD with a rate of 42% cases, thus we considered this terminal node as the reference. As shown in Table 2, individuals carrying the combination of *CASP3* rs2720378 GC or CC genotype, *IL-2* rs2069762 AC or CC genotype and *STIM1* rs1561876 AA genotype exhibited the highest risk for KD (OR = 2.12, 95% CI = 1.46–3.07, $P < 0.001$).

Afterwards, we attempted to test pairwise interactions between the three SNPs (rs2720378, rs2069762 and rs1561876) by the usual logistic regression analysis. However, we did not find any positive result on the interaction terms, regardless of multiplicative interaction or additive interaction (Supplementary Table S2 online).

**Cumulative effect analysis.** A cumulative effect of the 3 SNPs identified in the CART analysis was evaluated by LR analysis with the rs2720378 C allele, rs2069762 C allele and rs1561876 A allele as risk alleles. Participants were categorized into 3 groups based on the number of risk alleles they harbored ($0 \sim 2$, $3 \sim 4$, $5 \sim 6$ as individual group respectively, for the number of 0 risk allele subgroup was considerably small) and the $0 \sim 2$ risk alleles subgroup was regarded as the reference group. Both of the subgroups with $3 \sim 4$ risk alleles and $5 \sim 6$ risk alleles showed marginally significant or significant associations with increased KD risk (OR = 1.32, 95% CI = 1.00–1.74, $P = 0.052$; OR = 2.93, 95% CI = 1.66–5.17, $P < 0.001$), and the Cochran-Armitage trend test indicated a significant dose effect with the ORs being increased with increasing numbers of risk alleles ($P_{trend} < 0.001$, Table 3).

## Discussion

KD is an immune-mediated multi-systemic vasculitis and the most widely proposed consensus is that it results from an unknown infection trigger in part of genetically susceptible hosts[19,30]. A number of genes have been reported to have significant associations with the susceptibility to KD in different populations, such as *CD40L*[31], *HLA-E*[32], *FAM167A-BLK*[33], *FCGR2A*[34], through candidate gene approaches or genome-wide association studies. However, most of

these genes were discrete no matter which strategies were applied. Transforming growth factor-beta signaling pathway, as the only pathway studied on KD hitherto, has been proved to play an important role in KD pathogenesis and genetic variants in three genes in the pathway (*TGFB2*, *TGFBR2*, *SMAD3*) influence KD susceptibility, coronary artery aneurysms, aortic root dilatation and intravenous immunoglobulin treatment response[35].

To the best of our knowledge, this is probably the first study on the single- and multiple-risk of genetic variants in the $Ca^{2+}$/NFAT signaling pathway and KD risk in a Chinese population. In our study, the candidate SNPs were predicted to be functional by bioinformatics analyses or identified by previous studies. We first demonstrated that variants of *CASP3* rs2720378 and *IL-2* rs2069762 were nominally associated with increased risk of KD. The following CART analysis revealed the prediction value of gene-gene interactions among *CASP3* rs2720378, *IL-2* rs2069762 and *STIM1* rs1561876 variants on KD risk, and the cumulative analysis further indicated their synergetic effect on KD risk.

$Ca^{2+}$/NFAT signaling pathway plays a crucial role in the immune cell functions, including cell proliferation, cytokine gene expression, differentiation, and cell death[36]. Accumulated evidence has indicated that any abnormal expression or anergy caused by mutations of any component in $Ca^{2+}$/NFAT signaling pathway could be closely related to the change of signal intensity or transmission, thus contributing to the pathogenesis or progression of diseases, especially the immune-related diseases[37–42]. As far as our current findings were concerned, one of the positive results on SNP rs2720378 was also observed in one previous study conducted by Onouchi et al.[19]. The investigators explored the biological effect of *CASP3* rs113420705, and found that the G > A substitution weakened the binding between *CASP3* and NFATc2, reduced *CASP3* transcription, and presumably would inhibit T-cell apoptosis. In addition, rs2720378 was in tight linkage disequilibrium (LD) with rs113420705 in Japanese ($r^2 = 0.88$). Accordingly, we could speculate that rs2720378 might not be the "real" causal variant but only a "proxy". However, as Onouchi et al. pointed out in the study, there remained a possibility that rs2720378 also affected CASP3 expression by other, unknown mechanisms. Using the *in silico* tool SNPinfo[43], rs2720378 was classified as affecting transcription-factor-binding site (TFBS) activity, with the moderate difference score between the two alleles, which might be a possible hint. On the other hand, SNP rs2069762, a single base change proximal to the *IL-2* promoter region ($-330A > C$), was

**Table 3 | Cumulative effect of the 3 SNPs (rs2720378 rs2069762 rs1561876) between KD patients and normal controls**

| Number of risk alleles | KD cases (%) | Controls (%) | OR | 95% CI | LR P* |
|---|---|---|---|---|---|
| 0–2 | 144 (35.38) | 216 (44.08) | Ref. | Ref. | |
| 3–4 | 222 (54.55) | 253 (51.63) | **1.32** | **1.00–1.74** | **0.052** |
| 5–6 | 41 (10.07) | 21 (4.29) | **2.93** | **1.66–5.17** | **<0.001** |
| Cochran-Armitage Trend Test | | | **$P < 0.001$** | | |

Abbreviations: KD, Kawasaki disease; OR, odds ratio; 95%CI, 95% confidence interval; LR, logistic regression
*All the P values were adjusted for gender. The positive results were in bold

also predicted as affecting TFBS activity, where the C allele had a significantly higher score than the A allele. What's more, it has been proved that individuals homozygous for the C allele of rs2069762 produced over three times the amount of IL-2 than their AA and AC counterparts after stimulation with anti-CD3/CD28[25], and then correlated with a series of inflammatory or immunological diseases, such as ulcerative colitis[40], multiple sclerosis[41], and childhood lymphoma[42]. Therefore, we considered that it was biologically plausible that SNP rs2069762 was also associated with increased KD risk. Referring to *STIM1* rs1561876, few studies have investigated its effect on KD risk, but a recent study has demonstrated that the G allele of rs1561876 was associated with lower intracellular 1-β-D-arabinofur-anosyl-CTP levels, inferior response after remission induction therapy, greater risk of relapse and poorer overall survival in acute myeloid leukemia patients receiving cytarabine and cladribine[44]. Moreover, this SNP was predicted as affecting TFBS activity as well as microRNA-binding site activity, which might provide a clue for further study in the future. The multiple-risk genetic variants detected by CART, *CASP3* rs2720378, *IL-2* rs2069762 and *STIM1* rs1561876, suggested a stronger combined effect of SNPs, which might evoke researchers to pay more attention to the gene-gene interactions contributing to KD risk if they really exist.

Although the association of *ITPKC* rs28493229 with KD susceptibility has been reported in patients of Japanese and American ethnicities[12,45], we failed to reproduce such findings in our current study, which was in accordance with a recent study on Chinese population from Sichuan province of China in 2012[14]. The relatively low statistical power caused by insufficient number of study subjects might be taken into account. Moreover, the significantly lower frequency of the risk-conferring C allele of rs28493229 in CHB (7%) compared with Japanese in Tokyo or Americans (both are 15%), which data was obtained from 1000 Genomes (http://www.1000genomes.org/), might simply reflect the ethnic variation.

Considering the vital importance of Ca$^{2+}$/NFAT pathway in pathogenesis of KD, exploring novel inhibitors against this pathway seems to be promising in treatment of KD. Recently, two clinical trials of cyclosporine A, which potently suppresses the activity of T cells as a calcineurin inhibitor, have indicated that cyclosporine A appeared to be a safe and effective approach for patients with refractory KD[46,47].

Certain limitations to this study should be noted. Firstly, the candidate genes we studied were the primary but not all the genes involved in the Ca$^{2+}$/NFAT pathway, and the screening strategy of candidate SNPs was not so rigorous that the SNPs we selected could not cover all the probably functional SNPs among these gene regions. Thus, it might be unable to fully reflect the genetic effects of this pathway on KD susceptibility and might reduce the effectiveness of our conclusions. Secondly, these findings were preliminary since the sample size was relatively small and they were not replicated in this current study. Thirdly, the lack of information about environment factors, such as family history, infection history, which might play roles in KD onset, limited our further investigation of gene-environment interactions. Fourthly, the significant associations we observed were merely on the statistical level, whereas we did not evaluate their real biological functions.

In summary, our study implicates the importance of the single- and multiple-risk variants in Ca$^{2+}$/NFAT signaling pathway in KD susceptibility, and provides a clue to detect the potential gene-gene interactions conferring KD risk. Further studies on more comprehensive SNPs, different ethnicities and larger sample sizes are warranted, and follow-up studies are also needed to further identify the true causal variants to reveal the biological mechanism of genetic etiology.

## Methods

**Study subjects.** This study contained 428 children diagnosed with KD, all of which were unrelated ethnic Han Chinese and were consecutively recruited between April 2009 and September 2012 from Children's Hospital, Zhejiang University School of Medicine, China. The diagnosis of KD was based on the 5th revised edition of the guidelines established by the Kawasaki Disease Research Committee in Japan in 2002 (http://kawasaki-disease.org/diagnostic/). The controls comprised 493 ethnically and gender-matched unrelated healthy Chinese children without any evidence of infection. They were recruited from the same hospital at the time of a routine physical examination.

This study was approved by the ethics committees of the Children's Hospital, Zhejiang University School of Medicine, and the methods were carried out in accordance with the approved guidelines. Informed consent was obtained from all participants or parents/caregivers of all the subjects who were studied.

**Identification of candidate SNPs and genotyping.** The candidate genes involved in the Ca$^{2+}$/NFAT pathway were selected based on recent findings on genetic variants of KD, previous association studies and biological evidences, including 7 genes (*ITPKC*, *CASP3*, *STIM1*, *STIM2*, *NFATc1*, *NFATc2* and *IL-2*). The screening procedure of SNPs which were predicted to be functional in these genes was described as follows. First, we extracted the ranges of the physical positions covering the 7 genes and their 2 Kb range up- and downstream from the HapMap database (http://hapmap.ncbi.nlm.nih.gov/, HapMap Data Rel 24/phaseII Nov08, on NCBI B36 assembly, dbSNP b126), CHB. Second, we placed the position of each gene into an integrated bioinformatics tool "SNPinfo"[43] (http://snpinfo.niehs.nih.gov/snpinfo/snpfunc.htm) and retrieved a set of SNPs with possible functions. Third, we filtered these SNPs by MAF of CHB >10%, and got a total of 52 SNPs in the 7 candidate genes. As shown in Supplementary Table S3 online, the detailed information of all the SNPs were listed, including the possible functions of allelic difference which the SNPinfo tool can predict (marked with a Y letter), such as TFBS, splicing site, microRNA-binding site, nonsynonymous SNP, polymorphism phenotyping. Forth, we tested the LD among these SNPs by Haploview v4.2[48] (Supplementary Fig. S1 online), and SNPs in strong LD with each other ($r^2 \geq 0.80$) were considered redundant and only one was reserved. As a result, a total of 32 SNPs were remained for further selection (Supplementary Table S4 online). Since a few of these SNPs, such as *ITPKC* rs7251246, *STIM1* rs12806698, did not present in the results of Haploview, we rechecked the LD of these 32 SNPs using the online database of 1000 Genomes (http://www.1000genomes.org/), and found that several pairs of SNPs were in strong LD with each other (Supplementary Table S5 online). Therefore, 27 SNPs were obtained applying the similar filtering procedure (Supplementary Table S6 online). Next, we set priority of SNPs predicted to harbor multiple functional motifs if there also existed SNPs with single prediction of functional effect in the same gene (Supplementary Table S7 online). Afterwards we deleted five SNPs (rs9962479, rs4811191, rs2581732, rs9518 and rs4647601), among the genes which harbored more than one probably functional SNP.

Moreover, we added *CASP3* rs113420705 and *ITPKC* rs28493229, which have been previously identified as biologically functional and associated with KD risk[12,19]. Thus, a total of 16 SNPs were collected for this study (Supplementary Table S1 online).

The allelic differences of effect on DNA motif which the SNPinfo could predict were shown in Supplementary Table S8 online. For different possible functions, the prediction scores were calculated by different methods, such as Match[49], miRanda[50], ESEfind[51] and PolyPhen[52], and if there were various pairs of scores for two alleles of a given SNP, we chose the one which has the maximal difference score between the two alleles. The only exception was rs2290692, with the same score between the two alleles. Since it has been found associated with KD susceptibility in a Chinese population in a previous study[14], we reserved it for further study.

Genomic DNA was extracted from 2 mL peripheral blood sample that was collected from each participant at recruitment, using the RelaxGene Blood System DP319-02 (Tiangen, Beijing, China) by reference to the manufacturer's instructions. The concentration and the optical density of DNA were confirmed by NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham, Massachusetts, USA). All the SNPs were genotyped using TaqMan OpenArray Genotyping Assay System (Applied Biosystems, Foster City, CA, USA), the main procedure of which was described as follows. The final reaction volume per well was 4 μL for a 384-well plate, including TaqMan OpenArray Mater Mix, 2 × (2 μL) and normalized DNA sample (2 μL), the former of which was customized by Applied Biosystems. We applied the OpenArray AccuFill System (Applied Biosystems, Foster City, CA, USA) to load the samples into OpenArray plates (32 × 96 format) automatically, then we placed the loaded plates into cycling case, UV cured for 90 seconds. Next, we loaded the plates into the thermal cycler with the initial temperature of 93°C for 10 minutes, followed by 50 cycles of 95°C for 45 seconds, 94°C for 13 seconds and 53°C for 134 seconds, and after that, the temperature dropped to 25°C for 2 minutes and was maintained at 4°C. Finally, allelic discrimination plate read and analysis were performed using TaqMan OpenArray System, followed the manufacturer's protocol.

Notably, genotyping was performed without knowing the status of the participants. A 5% masked, random samples were genotyped twice for quality control, and the reproducibility was 100%.

**Statistical methods.** SNPs with genotyping call rates <95% or those that deviated from the HWE in controls ($P < 0.05$) were excluded. The HWE for genotypes was assessed by a goodness-of-fit $\chi^2$ test in the control group. The distribution differences in gender and genotype frequencies between cases and controls were examined by Pearson's $\chi^2$ test or Fisher's exact test, when appropriate. Unconditional LR analysis was then applied to estimate ORs and their 95% CIs for the effects of the SNPs on KD susceptibility, under assumption that variant alleles were the risk alleles. In order to increase the statistical power, the most likely inheritance model for each SNP was

selected instead of three models calculated simultaneously. The statistical power to detect the effects of the SNPs was calculated by Power v3.0.0[53,54], and for example, for SNPs with minor allele frequency of 0.070, 0.244 and 0.500, we calculated that the power for our sample size to detect an odds ratio of 1.50 was 0.40, 0.79 and 0.86, respectively.

The potential gene-gene interactions were examined by CART, which is a non-parametric technique and does not require assumptions about the distribution of the data. A CART tree is constructed by splitting data recursively into binary subsamples, beginning with the root node that contains all the learning sample, and leading to the formation of daughter nodes (nodes that can be split further) and terminal nodes (nodes that cannot be split any further). Gini criteria are used to achieve a high degree of homogeneity in the terminal nodes or subgroups. After a large tree is grown, which tends to be complex and difficult to interpret, a pruning procedure is performed to avoid overfitting the model. In the final stage, the optimal tree is selected based on the lowest misclassification error rate, which can be assessed based on cross validation[55]. Subgroups of individuals with differential risk associations with KD were identified in the different terminal nodes of the tree, indicating potential interactions. Finally, the risk of these subgroups was evaluated by LR analysis with the least percentage of cases as the reference and ORs and 95% CIs adjusted for gender[28]. After CART analysis, multiple risk of the variants potentially interacted with each other was evaluated by a cumulative effect analysis.

All of the above statistical analyses were conducted by SPSS software v13.0 (SPSS, Inc., Chicago, IL) and all $P$ values were two-tailed with a statistically significant level at 0.05.

1. Kawasaki, T., Kosaki, F., Okawa, S., Shigematsu, I. & Yanagawa, H. A new infantile acute febrile mucocutaneous lymph node syndrome (MLNS) prevailing in Japan. *Pediatrics* **54**, 271–276 (1974).
2. Nakamura, Y. *et al.* Epidemiologic features of Kawasaki disease in Japan: Results of the 2009–2010 nationwide survey. *J. Epidemiol.* **22**, 216–221 (2012).
3. Park, Y. W. *et al.* Epidemiological features of Kawasaki disease in Korea, 2006–2008. *Pediatr. Int.* **53**, 36–39 (2011).
4. Lue, H. C. *et al.* Epidemiological features of Kawasaki disease in Taiwan, 1976–2007: Results of five nationwide questionnaire hospital surveys. *Pediatr. Neonatol.* DOI: 10.1016/j.pedneo.2013.07.010 (2013).
5. Newburger, J. W. *et al.* Diagnosis, treatment, and long-term management of Kawasaki disease: A statement for health professionals from the committee on rheumatic fever, endocarditis, and Kawasaki disease, council on cardiovascular disease in the young, American Heart Association. *Circulation* **110**, 2747–2771 (2004).
6. Burns, J. C. & Glodé, M. P. Kawasaki syndrome. *Lancet* **364**, 533–544 (2004).
7. Kato, H., Koike, S., Yamamoto, M., Ito, Y. & Yano, E. Coronary aneurysms in infants and young children with acute febrile mucocutaneous lymph node syndrome. *J. Pediatr.* **86**, 892–898 (1975).
8. Kato, H. *et al.* Long-term consequences of Kawasaki disease. A 10- to 21-year follow-up study of 594 patients. *Circulation* **94**, 1379–1385 (1996).
9. Fujita, Y. *et al.* Kawasaki disease in families. *Pediatrics* **84**, 666–669 (1989).
10. Uehara, R., Yashiro, M., Nakamura, Y. & Yanagawa, H. Kawasaki disease in parents and children. *Acta Paediatr.* **92**, 694–697 (2003).
11. Onouchi, Y. *et al.* A genomewide linkage analysis of Kawasaki disease: Evidence for linkage to chromosome 12. *J. Hum. Genet.* **52**, 179–190 (2007).
12. Onouchi, Y. *et al.* ITPKC functional polymorphism associated with Kawasaki disease susceptibility and formation of coronary artery aneurysms. *Nat. Genet.* **40**, 35–42 (2008).
13. Lou, J. *et al.* A functional polymorphism, rs28493229, in ITPKC and risk of Kawasaki disease: An integrated meta-analysis. *Mol. Biol. Rep.* **39**, 11137–11144 (2012).
14. Peng, Q. *et al.* Single-nucleotide polymorphism rs2290692 in the 3′UTR of ITPKC associated with susceptibility to Kawasaki disease in a Han Chinese population. *Pediatr. Cardiol.* **33**, 1046–1053 (2012).
15. Brown, T. J. *et al.* CD8 T lymphocytes and macrophages infiltrate coronary artery aneurysms in acute Kawasaki disease. *J. Infect. Dis.* **184**, 940–943 (2001).
16. Terai, M. *et al.* Imbalance among T-cell subsets in patients with coronary arterial aneurysms in Kawasaki disease. *Am. J. Cardiol.* **60**, 555–559 (1987).
17. Rowley, A. H. Kawasaki disease: Novel insights into etiology and genetic susceptibility. *Annu. Rev. Med.* **62**, 69–77 (2011).
18. Wu, W. *et al.* Proteolytic regulation of nuclear factor of activated T (NFAT) c2 cells and NFAT activity by caspase-3. *J. Biol. Chem.* **281**, 10682–10690 (2006).
19. Onouchi, Y. *et al.* Common variants in CASP3 confer susceptibility to Kawasaki disease. *Hum. Mol. Genet.* **19**, 2898–2906 (2010).
20. Roos, J. *et al.* STIM1, an essential and conserved component of store-operated Ca2+ channel function. *J. Cell Biol.* **169**, 435–445 (2005).
21. Brandman, O., Liou, J., Park, W. S. & Meyer, T. STIM2 is a feedback regulator that stabilizes basal cytosolic and endoplasmic reticulum Ca2+ levels. *Cell* **131**, 1327–1339 (2007).
22. Oh-Hora, M. *et al.* Dual functions for the endoplasmic reticulum calcium sensors STIM1 and STIM2 in T cell activation and tolerance. *Nat. Immunol.* **9**, 432–443 (2008).
23. Peng, S. L., Gerth, A. J., Ranger, A. M. & Glimcher, L. H. NFATc1 and NFATc2 together control both T and B cell activation and differentiation. *Immunity* **14**, 13–20 (2001).
24. Macian, F. NFAT proteins: Key regulators of T-cell development and function. *Nat. Rev. Immunol.* **5**, 472–484 (2005).
25. Hoffmann, S. C. *et al.* Association of cytokine polymorphic inheritance and in vitro cytokine production in anti-CD3/CD28-stimulated peripheral blood lymphocytes. *Transplantation* **72**, 1444–1450 (2001).
26. Peduzzi, P., Concato, J., Kemper, E., Holford, T. R. & Feinstein, A. R. A simulation study of the number of events per variable in logistic regression analysis. *J. Clin. Epidemiol.* **49**, 1373–1379 (1996).
27. Concato, J., Feinstein, A. R. & Holford, T. R. The risk of determining risk with multivariable models. *Ann. Intern. Med.* **118**, 201–210 (1993).
28. Liu, L. *et al.* Association of candidate genetic variations with gastric cardia adenocarcinoma in Chinese population: A multiple interaction analysis. *Carcinogenesis* **32**, 336–342 (2011).
29. Garcia-Magarinos, M., Lopez-de-Ullibarri, I., Cao, R. & Salas, A. Evaluating the ability of tree-based methods and logistic regression for the detection of SNP-SNP interaction. *Ann. Hum. Genet.* **73**, 360–369 (2009).
30. Lee, K. Y., Han, J. W. & Lee, J. S. Kawasaki disease may be a hyperimmune reaction of genetically susceptible children to variants of normal environmental flora. *Med. Hypotheses* **69**, 642–651 (2007).
31. Onouchi, Y. *et al.* CD40 ligand gene and Kawasaki disease. *Eur. J. Hum. Genet.* **12**, 1062–1068 (2004).
32. Lin, Y. J. *et al.* HLA-E gene polymorphism associated with susceptibility to Kawasaki disease and formation of coronary artery aneurysms. *Arthritis Rheum.* **60**, 604–610 (2009).
33. Wang, W. *et al.* 8p22-23-rs2254546 as a Susceptibility Locus for Kawasaki Disease: a Case-control Study and a Meta-analysis. *Sci. Rep.* **4**, 4247; DOI: 10.1038/srep04247 (2014).
34. Khor, C. C. *et al.* Genome-wide association study identifies FCGR2A as a susceptibility locus for Kawasaki disease. *Nat. Genet.* **43**, 1241–1246 (2011).
35. Shimizu, C. *et al.* Transforming growth factor-beta signaling pathway in patients with Kawasaki disease. *Circ. Cardiovasc. Genet.* **4**, 16–25 (2011).
36. Srikanth, S. & Gwack, Y. Orai1-NFAT signalling pathway triggered by T cell receptor stimulation. *Mol. Cells* **35**, 182–194 (2013).
37. Picard, C. *et al.* STIM1 mutation associated with a syndrome of immunodeficiency and autoimmunity. *N. Engl. J. Med.* **360**, 1971–1980 (2009).
38. Byun, M. *et al.* Whole-exome sequencing-based discovery of STIM1 deficiency in a child with fatal classic Kaposi sarcoma. *J. Exp. Med.* **207**, 2307–2312 (2010).
39. Shimada, M. *et al.* An approach based on a genome-wide association study reveals candidate loci for narcolepsy. *Hum. Genet.* **128**, 433–441 (2010).
40. Shi, J. *et al.* Haplotype-based analysis of ulcerative colitis risk loci identifies both IL2 and IL21 as susceptibility genes in Han Chinese. *Inflamm. Bowel Dis.* **17**, 2472–2479 (2011).
41. Matesanz, F. *et al.* Allelic expression and interleukin-2 polymorphisms in multiple sclerosis. *J. Neuroimmunol.* **119**, 101–105 (2001).
42. Song, N. *et al.* Genetic variants in interleukin-2 and risk of lymphoma among children in Korea. *Asian Pac. J. Cancer Prev.* **13**, 621–623 (2012).
43. Xu, Z. & Taylor, J. A. SNPinfo: Integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res.* **37**, W600–605 (2009).
44. Cao, X. *et al.* RRM1 and RRM2 pharmacogenetics: Association with phenotypes in HapMap cell lines and acute myeloid leukemia patients. *Pharmacogenomics* **14**, 1449–1466 (2013).
45. Onouchi, Y. *et al.* ITPKC and CASP3 polymorphisms and risks for IVIG unresponsiveness and coronary artery lesion formation in Kawasaki disease. *Pharmacogenomics J.* **13**, 52–59 (2013).
46. Suzuki, H. *et al.* Cyclosporin A treatment for Kawasaki disease refractory to initial and additional intravenous immunoglobulin. *Pediatr. Infect. Dis. J.* **30**, 871–876 (2011).
47. Tremoulet, A. H. *et al.* Calcineurin inhibitor treatment of intravenous immunoglobulin-resistant Kawasaki disease. *J. Pediatr.* **161**, 506–512.e1 (2012).
48. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
49. Kel, A. E. *et al.* Match: A tool for searching transcription factor binding sites in DNA sequences. *Nucleic Acids Res.* **31**, 3576–3579 (2003).
50. John, B. *et al.* Human MicroRNA targets. *PLoS Biol.* **2**, e363 (2004).
51. Cartegni, L., Wang, J., Zhu, Z., Zhang, M. Q. & Krainer, A. R. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* **31**, 3568–3571 (2003).
52. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
53. Lubin, J. H. & Gail, M. H. On power and sample size for studying features of the relative odds of disease. *Am. J. Epidemiol.* **131**, 552–566 (1990).
54. Garcia-Closas, M. & Lubin, J. H. Power and sample size calculations in case-control studies of gene-environment interactions: Comments on different approaches. *Am. J. Epidemiol.* **149**, 689–692 (1999).
55. Kitsantas, P. & Wu, H. Body mass index, smoking, age and cancer mortality among women: A classification tree analysis. *J. Obstet. Gynaecol. Res.* **39**, 1330–1338 (2013).

## Acknowledgments

## Author contributions

Conceived and designed the study strategy: X.M., F.G. Designed the experiment: W.W., J.L. Recruited the participants and collected their information and blood samples: J.L., X.L. Performed the experiments: J.L., R.Z., N.S. Analysis the data: W.W., J.L., Y.Q. Collected and checked the information of eligible studies included in meta-analysis: W.W., J.L. Statistical analyses: W.W., N.S., Y.W., Q.Z. Reference collection and data management: J.D., J.K., L.Z. Wrote the manuscript: W.W., J.L. Prepared the tables and figures: W.W., X.L. All authors reviewed the manuscript.

## Additional information

**Supplementary information** accompanies this paper at http://www.nature.com/scientificreports

**Competing financial interests**: The authors declare no competing financial interests.

**How to cite this article**: Wang, W. *et al.* The Roles of $Ca^{2+}$/NFAT Signaling Genes in Kawasaki Disease: Single- and Multiple-Risk Genetic Variants. *Sci. Rep.* **4**, 5208; DOI:10.1038/srep05208 (2014).